

# Local tomography and the Jordan structure of quantum theory

Howard Barnum<sup>\*</sup> and Alexander Wilce<sup>†</sup>

February 22, 2012

## Abstract

Using a result of H. Hanche-Olsen, we show that (subject to fairly natural constraints on what constitutes a system, and on what constitutes a composite system), orthodox finite-dimensional complex quantum mechanics with superselection rules is the only non-signaling probabilistic theory in which (i) individual systems are Jordan algebras (equivalently, their cones of unnormalized states are homogeneous and self-dual), (ii) composites are locally tomographic (meaning that states are determined by the joint probabilities they assign to measurement outcomes on the component systems) and (iii) at least one system has the structure of a qubit. Using this result, we also characterize finite dimensional quantum theory among probabilistic theories having the structure of a dagger-monoidal category.

## 1 Introduction and background

One of the oldest foundational problems besetting quantum mechanics is to provide a clear motivation for its probabilistic apparatus — in particular, for the representation of observables of a quantum system by the self-adjoint elements of a  $C^*$  algebra. *Why* should outcomes of measurements give rise to anything so nicely structured as a  $C^*$ -algebra — or any algebra at all, for that matter? In particular, what operational meaning can we give to the product of two non-commuting observables, when these cannot simultaneously be measured, and when, indeed, this product is not self-adjoint?

### 1.1 Jordan algebras

In an early attempt to address this question, Pascual Jordan [21] proposed in 1932 that the observables associated with a finite-dimensional physical system should constitute what is now called a formally real Jordan algebra. A Jordan algebra is a finite-dimensional real vector space  $\mathbf{E}$  equipped with a commutative bilinear operation  $\bullet : \mathbf{E} \times \mathbf{E} \rightarrow \mathbf{E}$  satisfying the *Jordan identity*

$$a^2 \bullet (a \bullet b) = a \bullet (a^2 \bullet b)$$

for all  $a, b \in \mathbf{E}$  (where  $a^2 := a \bullet a$ ). Jordan algebras are naturally equipped with a bilinear trace form, which induces a symmetric, nondegenerate bilinear form  $(a, b) \mapsto \langle a, b \rangle := \text{tr}(a \bullet b)$ . If this is an inner product (that is, positive-definite), one calls  $\mathbf{E}$  *Euclidean*. In finite dimensions, this is

---

<sup>\*</sup>Department of Physics and Astronomy, University of New Mexico; [hnbarnum@aol.com](mailto:hnbarnum@aol.com), [hnbarnum@unm.edu](mailto:hnbarnum@unm.edu)

<sup>†</sup>Department of Mathematics, Susquehanna University; [wilce@susqu.edu](mailto:wilce@susqu.edu)

equivalent to Jordan's condition of *formal reality*: that  $a^2 + b^2 = 0 \Rightarrow x = y = 0$ . Two years later, Jordan, von Neumann, and Wigner [22] classified such algebras as being either (i) self-adjoint parts of (real, complex or quaternionic) matrix algebras, under the anti-commutator  $x \bullet y = (xy + yx)/2$ , (ii) so-called spin factors, or (iii) the self-adjoint part of the 3-by-3 matrix algebra over the octonions, or direct sums of these. Thus, the assumption that the space of observables of a physical system is a Euclidean Jordan algebra does bring one very close to finite-dimensional quantum mechanics.

## 1.2 General probabilistic theories

An early objection to the Jordan-algebraic approach is that it does not generalize easily to the infinite-dimensional setting required for full-blown quantum mechanics. However, in recent years, with the growing importance of quantum information theory, finite-dimensional quantum theory is coming to be viewed as an important subject in its own right, and is even regarded (in some quarters) as being possibly more fundamental than more traditional, infinite-dimensional QM. Moreover, the introduction of the notion of JB-algebra has turned out to provide a fairly satisfactory generalization to infinite dimension.

Leaving this issue to one side, another, and more basic, objection is that the Jordan product has no clearer an operational interpretation than the  $C^*$ -algebraic product. Later work has tended to start with a much more general (but conceptually much more transparent) framework [26, 20, 5], in which a physical — or, more broadly, probabilistic — system is represented by an *order-unit space*. This is an ordered real vector space  $\mathbf{E}$ , with positive cone  $\mathbf{E}_+$ , equipped with a distinguished element  $u \in \mathbf{E}_+$ , called the *order unit*, such that for every  $a \in \mathbf{E}_+$ ,  $ta \leq u$  for some  $t > 0$ . Possible measurement outcomes associated with the system are identified with *effects*, that is, vectors  $a \in \mathbf{E}_+$  with  $a \leq u$ . States of the system are identified with positive linear functionals  $\alpha : \mathbf{E} \rightarrow \mathbb{R}$  with  $\rho(u) = 1$ . If  $a$  is an effect, then  $\alpha(a)$  is understood to be the probability that  $a$  will occur (if measured) when the state  $\alpha$  obtains. Physical processes acting on a system, or between two systems, can then be represented very naturally by positive linear mappings between the associated ordered linear spaces, and one can define a *probabilistic theory* to be a category of such spaces and mappings.

Within this very general setting (which we review in greater detail in Section 2), one can hope to find illuminating characterizations of quantum theory, that is, theorems that single out QM — particularly, complex QM — as the unique probabilistic theory satisfying one or more reasonable constraints. This was the goal, explicit or tacit, of a great deal of foundational work in quantum theory from roughly the 1950s to the late 1970s [26, 25]. With the emergence of quantum information theory, this project has enjoyed a strong revival, with a distinctive focus on finite-dimensional systems, and an emphasis on composite systems [19, 30, 17, 14, 12, 27]. The cited papers all come close to, or indeed succeed in, deriving finite dimensional QM from simple axioms. However, many [19, 30, 12, 27] make use of a strong uniformity principle, namely, that all systems having the same information-carrying capacity (as variously defined) are isomorphic; others place strong constraints on the representation of sub-systems [19, 14]. We hope to avoid both kinds of assumptions.

## 1.3 Homogeneity and self-duality

A different approach, which we have pursued in [7, 8, 37, 38], is to exploit the classical correspondence between Jordan algebras and homogeneous self-dual cones. This is reviewed in more detail below, but, briefly: The positive cone  $\mathbf{E}_+$  of an ordered vector space  $\mathbf{E}$  is *homogeneous* iff the group of order-automorphisms<sup>1</sup> of  $\mathbf{E}$  acts transitively on the *interior* of  $\mathbf{E}_+$ , and *self-dual* iff there exists an

<sup>1</sup>That is, positive linear bijections having positive inverses

inner product on  $\mathbf{E}$  such that

$$\mathbf{E}_+ = \mathbf{E}^+ := \{a \in \mathbf{E} \mid \langle a, b \rangle \geq 0 \ \forall b \in \mathbf{E}_+\}.$$

**Theorem (Koecher [24], Vinberg [34]):** *Let  $\mathbf{E}$  be a finite-dimensional order-unit space with a homogeneous, self-dual (HSD) cone  $\mathbf{E}_+$ . Then there exists a unique bilinear operation  $\bullet : \mathbf{E} \times \mathbf{E} \rightarrow \mathbf{E}$  making  $\mathbf{E}$  into a Euclidean Jordan algebra with unit  $u$  and cone of squares equal to  $\mathbf{E}_+$ .*

In [7], we observed that a simple purification or dilation principle is enough to guarantee that the cone of states of a physical system is homogeneous and *weakly* self-dual, i.e.,  $\mathbf{E}_+^* \simeq \mathbf{E}_+$ . However, the distinction between weak self-duality and self-duality is significant, so this result still leaves us with two questions: first, why the state cone ought to be self-dual, and, secondly, how to rule out, or to make room for, the various alternatives to complex QM allowed by the Jordan-von-Neumann-Wigner classification.

In this paper, we bracket the first question (to which several possible answers have been suggested; see [7, 8, 28, 37, 38]) and concentrate on the second. We consider a probabilistic theory in which (i) individual systems are represented by homogeneous, self-dual models — equivalently, by formally real Jordan algebras — and ask when these must in fact be standard *quantum* models, i.e, the self-adjoint parts of complex matrix algebras.

## 1.4 Composites of homogeneous, self-dual systems

As it happens, a *nearly* off-the-shelf answer is available. It has been known at least since [3] that complex QM is distinguished from its real analogue by a property called *local tomography*, which requires that the joint state of a composite system be completely determined by the joint probabilities assigned to observables on the two component systems.<sup>2</sup> In [18], H. Hanche Olsen made a similar point regarding Jordan algebras:

**Theorem (Hanche-Olsen, [18]):** *Let  $\mathbf{E}_2$  be the Jordan algebra of hermitian  $2 \times 2$  complex matrices, i.e., the Jordan algebra corresponding to a single qubit. Let  $\mathbf{E}$  be any JB algebra (in finite-dimensions, the same thing as a Euclidean Jordan algebra), and suppose that the vector space  $\mathbf{E} \otimes \mathbf{E}_2$  carries a Jordan product satisfying*

$$(a \otimes \mathbf{1}) \bullet (b \otimes v) = (a \bullet b) \otimes v \text{ and } (\mathbf{1} \otimes v) \bullet (a \otimes w) = a \otimes (v \bullet w). \quad (1)$$

*for all  $a, b \in \mathbf{E}$  and all  $v, w \in \mathbf{E}_2$ . Then  $\mathbf{E}$  is the Hermitian part of a  $C^*$ -algebra.*

Of course, absent a direct physical or operational interpretation of the Jordan product, Hanche-Olsen's condition (1) calls for some further motivation. In Section 4, we show that in the context of composites of probabilistic models, (1) follows from *local tomography* — the condition that the joint state of a composite system is determined by the joint probabilities it assigns to outcomes of measurements on the two component systems — plus the condition that the self-dualizing inner product on a composite system can be chosen so as to factor into a product of self-dualizing inner products for the component systems. We call a theory satisfying the latter condition *factorizably self-dual*. By a *factorizably HSD theory*, we mean a probabilistic theory in which every system is homogeneous and factorizably self-dual. Hanche-Olsen's result then yields

<sup>2</sup>It is also observed in [3] that in the quaternionic analogue of complex quantum theory, the most obvious candidate for the state space of a composite of  $m$ -dimensional and  $n$ -dimensional quaternionic systems, namely the  $mn \times mn$  dimensional positive semidefinite (PSD) quaternionic matrices, suffers from difficulties in even identifying the product effects necessary to a locally tomographic composite—indeed, its dimension is smaller than the product of the dimensions of the spaces spanned by the  $m \times m$  and by the  $n \times n$  quaternionic positive semi-definite matrices.

**Proposition 1.** *Let  $\mathcal{C}$  be any factorizably HSD probabilistic theory in which (i) every pair of systems  $A$  and  $B$  admit a locally-tomographic composite system,  $AB$ , still belonging to  $\mathcal{C}$ , and (ii) there exists a qubit. Then all systems in  $\mathcal{C}$  are self-adjoint parts of complex matrix algebras.*

The factorizability assumption can itself be further motivated. In particular, it is automatically satisfied given two very weak and natural conditions, namely, that each component system support a *uniform* (or *maximally mixed*) state, and that every basic measurement outcome have probability one in *some* state. Given these assumptions, finite-dimensional QM is completely characterized among finite-dimensional HSD theories by conditions (i) and (ii) above.

Proposition 1 has an important consequence for the categorical formulation of quantum theory in terms of dagger-monoidal categories [1, 4, 33]. Let  $\mathcal{C}$  be a dagger-monoidal category whose objects are order-unit spaces, with the set of morphisms between any two objects being a cone of positive linear mappings between these spaces, with tensor unit  $I \simeq \mathbb{R}$ . If  $\mathcal{C}(I, A) \simeq A$ , then each object  $A$  is equipped with a canonical bilinear form, namely  $\langle a, b \rangle = a^\dagger \circ b$ , which factors on tensor products. If this is an inner product, and the group of invertible elements of  $\mathcal{C}(A, A)$  acts homogeneously on  $A$ , then the positive cone of  $A$  is self-dual. Thus, if tensor products in  $\mathcal{C}$  are locally tomographic, non-signaling composites, and if  $\mathcal{C}$  contains a qubit, then every order-unit space  $A \in \mathcal{C}$  is the hermitian part of a  $C^*$  algebra.

The balance of this note supplies the proof of Proposition 1, along with enough technical background to make the exposition self-contained. In Section 2, we give a more detailed sketch of the general probabilistic framework described above, and discuss the structure of models associated with formally real Jordan algebras. In Section 3, after discussing composite systems in general, we study locally tomographic composites of Jordan-algebraic systems, and prove Proposition 1. In Section 4, we reconsider these ideas in the context of a dagger-monoidal category of probabilistic models. Section 5 offers a few concluding remarks, questions, and speculative suggestions.

## 2 Probabilistic Models and Theories

In this section we provide a quick review of the framework for generalized probability theory that we shall use. This is fairly standard, with a history going back ultimately to the work of Mackey in the 1950s. The precise machinery we use combines ideas borrowed from [13, 20, 16], here specialized to finite-dimensional systems.

### 2.1 States, Effects and Processes

In its very simplest formulation, classical probability theory concerns an “experiment” — a single, discrete set  $E$  of mutually exclusive possible outcomes, and probability weights thereon. A particularly simple (and, conceptually, very conservative) generalization of classical probability theory begins with the idea that one may be faced with a choice of experiments.

**Definition 2.** A *test space* is a family  $\mathfrak{A}$  of non-empty sets, called *tests*, construed as the outcome-sets associated with various experiments, measurements, or other operations. The *outcome space* of  $\mathfrak{A}$  is the set  $X := \bigcup \mathfrak{A}$  of all outcomes arising from any test  $E \in \mathfrak{A}$ . A *state*, or *probability weight*, on  $\mathfrak{A}$  is a mapping  $\alpha : X \rightarrow [0, 1]$  summing to unity on each  $E \in \mathfrak{A}$  — in other words,  $\alpha$  is a simultaneous (and non-contextual) assignment of a probability weight to each test.

**Examples:** (i) A discrete classical test space is one of the form  $\{E\}$ , that is, one that contains only a single test. (ii) One can also consider the test space consisting of finite (respectively, countable)

partitions of a measurable space  $S$  by measurable subsets; in this case, the states correspond exactly to finitely additive (respectively, countably additive) probability measures on  $S$ . (iii) The standard test space in quantum theory is the collection of maximal sets of pairwise orthogonal, rank-one projection operators on a Hilbert space  $\mathbf{H}$ . Gleason's Theorem tells us (for  $\dim(\mathbf{H}) > 2$ ) that all probability weights on this test space are implemented by density operators, according to the “Born rule”.

**States and Effects** Given a test space  $\mathfrak{A}$ , it is often reasonable to consider a restricted state space  $\Omega$ . (For instance, given a qubit, we typically restrict attention to those states given by density operators, rather than allowing the various discontinuous states that would otherwise be allowed by the very loose combinatorial structure of  $\mathbf{F}_2$ .) Plausibly,  $\Omega$  should be both convex and closed with respect to outcome-wise convergence — hence, compact as a subset of  $[0, 1]^X$ . It should also be rich enough to separate outcomes, in the sense that if  $x, y \in X$  and  $\alpha(x) = \alpha(y)$  for all  $\alpha \in \Omega$ , then  $x = y$ . We can now associate to every  $x \in X$  the corresponding evaluation functional  $\alpha \mapsto \alpha(x)$  in  $\mathbb{R}^\Omega$ . Let  $\mathbf{E}$  denote the span of  $X$  in  $\mathbb{R}^\Omega$ . We shall say that the pair  $(\mathfrak{A}, \Omega)$  is *finite-dimensional* iff  $\mathbf{E}$  is finite dimensional.

Now define a cone in  $\mathbf{E}$  by setting  $\mathbf{E}_+ = \{\sum_i t_i x_i \mid t_i \geq 0, x_i \in X\}$ . Let  $u$  denote the unit functional  $u(\alpha) \equiv 1$ ; then  $\sum_{x \in E} x = u$  for every test  $E \in \mathfrak{A}$ . In particular,  $x \leq u$  for every  $x \in X$ . It follows that  $u$  is an order-unit for  $\mathbf{E}$ . If  $\alpha \in \mathbf{E}^*$  is any normalized positive functional, i.e.  $\alpha(a) \geq 0$  for  $a \in \mathbf{E}_+$  and  $\alpha(u) = 1$ , then we obtain a state on  $\mathfrak{A}$  by restriction to  $X$ . The set of states arising in this way defines a compact convex set  $\hat{\Omega} \supseteq \Omega$ . Call  $\Omega$  *state-complete* iff  $\hat{\Omega} = \Omega$ . It is reasonable to assume, and we shall assume here, that **all state spaces are state-complete**. So for the remainder of the paper, “state” means “element of  $\hat{\Omega}$ ”.

**Processes** Any test space  $\mathfrak{A}$  is associated with a group of *symmetries*, i.e., bijections  $g : X \rightarrow X$  with  $gE \in \mathfrak{A} \leftrightarrow E \in \mathfrak{A}$  for all  $E \subseteq X$ . This group is compact in  $\mathbb{R}^X$ , and acts on  $\mathbf{E}$  by positive, unit-preserving linear automorphisms. Just as it may be reasonable to restrict the set of states, it may be desirable to consider a restricted set of symmetries. More generally, we may wish to identify a semigroup of “physical processes”. Such processes should surely map normalized states to possibly sub-normalized states, preserving convex combinations. Thus, we might represent a physical process by a positive mapping  $\phi : \mathbf{E}^* \rightarrow \mathbf{E}^*$ , with  $u(\phi(\alpha)) \leq u(\alpha)$  for all  $\alpha \in \mathbf{E}_+^*$ . We interpret  $u(\phi(\alpha))$  as the *probability* that  $\phi$  occurs when the initial state is  $\alpha$ .

If  $\phi : \mathbf{E}^* \rightarrow \mathbf{E}^*$  is a physical process, there will be a dual process  $\tau = \phi^* : \mathbf{E} \rightarrow \mathbf{E}$ , given by  $\phi^*(a) = a \circ \phi$  for any  $a \in \mathbf{E}$ . Operationally, to measure  $\phi^*(a)$  on a state  $\alpha$ , one first subjects the state  $\alpha$  to the process  $\phi$ , and then makes a measurement of the effect  $a$ . Note that  $\tau(u)(\alpha) = u(\tau^*(\alpha))$  is the probability that the process  $\tau^* = \phi$  occurs if the initial state is  $\alpha$ . In what follows, it will generally be more convenient to deal with these dual processes; accordingly, we'll broaden our usage and refer to these, also, as processes.

**Probabilistic Models and Theories** In view of the preceding discussion, the following language seems reasonable.

**Definition 3.** A finite-dimensional *probabilistic model* is a triple  $A = (\mathbf{E}(A), \mathfrak{A}(A), \mathcal{D}(A))$  consisting of

- (i) a finite-dimensional order-unit space  $(\mathbf{E}(A), u_A)$ ,
- (ii) a test space  $\mathfrak{A}$  consisting of observables on  $\mathbf{E}(A)$ , with outcome-set  $X = \bigcup \mathfrak{A}$  generating  $\mathbf{E}_+(A)$ , and
- (iii) a semigroup  $\mathcal{D}(A)$  of positive mappings  $\tau : \mathbf{E} \rightarrow \mathbf{E}$ , called *processes*, satisfying  $\tau(u) \leq u$ .

A *state* of the model is a normalized, positive linear functional  $\alpha : \mathbf{E}(A) \rightarrow \mathbb{R}$ .

Broadly speaking, a *probabilistic theory* is a class  $\mathcal{C}$  of such models. In particular, we can identify finite-dimensional quantum theory with the class of models in which  $\mathbf{E}$  is the set of hermitian elements of a complex matrix algebra  $\mathcal{A}$ , with the usual operator-theoretic ordering,  $u$  is the identity functional,  $\mathfrak{A}$  consists of maximal, pairwise orthogonal sets of projection operators, and  $\mathcal{D}$  is the semigroup of completely positive maps on  $\mathcal{A}$ .

**Reversible Processes** We shall say that a physical process  $\phi$ , or the dual process  $\tau = \phi^*$ , is *physically reversible* iff it is invertible as a linear mapping, with a positive inverse — that is,  $\phi$  is an *order-automorphism* of  $\mathbf{E}(A)^*$  — and  $\phi^{-1}$  is a positive multiple of a physical process — say,  $\phi^{-1} = c\phi_o$  for some process  $\phi_o$ . Operationally, this means that there is always some non-zero probability that  $\phi_o \circ \phi$  will return the system to its original state. Indeed, for any normalized state  $\alpha$ ,

$$\phi_o(\phi(\alpha))(u) = \phi_o(c\phi_o^{-1}(\alpha))(u) = c\alpha(u) = c,$$

so this probability — which is independent of the initial state  $\alpha$  — is exactly the factor  $c$ . Notice that  $\phi$  is reversible with probability one iff  $c = 1$ , i.e.,  $\phi^{-1}$  is a process.<sup>3</sup> This implies that  $\tau = \phi^*$  satisfies  $\tau(u) = u$ . Conversely, if  $\tau = \phi^*$  and  $\tau u = u$ , then  $\tau^{-1}u = u$ . Thus, if  $\tau^{-1} = c\tau_o$ , where  $\tau_o$  is a process, then, on states, then the probability of

$$(c\tau_o^*)(\alpha)(u) = c\alpha(\tau(u)) = c\alpha(u) = c.$$

Clearly, the set  $\mathcal{D}_1(A)$  of invertible processes forms a sub-semigroup of  $\mathcal{D}(A)$ , and generates a subgroup,  $\mathcal{G}(A)$ , of  $\text{Aut}(\mathbf{E}(A))$ , namely, the set of all multiples  $c\tau$  where  $\tau \in \mathcal{D}_1$  and  $c \in \mathbb{R}_+$ . Those processes reversible with probability 1 are exactly the invertible processes  $\tau \in \mathcal{D}(A)$  with  $\tau(u) = u$ , i.e., those in the stabilizer  $\mathcal{G}(A)_{u_A}$ .

## 2.2 The Jordan structure of an HSD model

Our proof of Proposition 1, given in Section 3, depends on the details of the construction of the Jordan product on an HSD order-unit space. In what follows, let  $(\mathbf{E}, u)$  be an HSD order-unit space. By this we mean a finite-dimensional order-unit space  $\mathbf{E}$ , the positive cone of which is homogeneous, and for which there *exists* an inner product making  $\mathbf{E}_+ = \mathbf{E}^+$ . We call such an inner product *self-dualizing*.<sup>4</sup> Let  $G$  be any closed subgroup of  $\text{Aut}(\mathbf{E})$ , acting transitively on the interior of  $\mathbf{E}_+$ . Then  $G$  is a Lie subgroup of  $GL(\mathbf{E})$ . Let  $\mathfrak{g}$  denote its Lie algebra, and let  $\mathfrak{g}_u$  denote the Lie algebra of the stabilizer  $G_u \leq G$  of the order-unit. The following formulation of the Koecher-Vinberg Theorem summarizes the construction of the Jordan product on  $\mathbf{E}$ .

**Theorem 4** (Koecher-Vinberg). *Let  $G$  be a closed, connected subgroup of  $\text{Aut}(\mathbf{E})$ , acting transitively on the interior of  $\mathbf{E}_+$ . Then*

- (a) *It is possible to choose a self-dualizing inner product on  $\mathbf{E}_+$  in such a way that  $G_u = G \cap \mathcal{O}(\mathbf{E})$  (where  $\mathcal{O}(\mathbf{E})$  is the orthogonal group with respect to the inner product);*
- (b) *If  $G = G^\dagger$  with respect to this inner product, then  $\mathfrak{g}_u = \{X \in \mathfrak{g} | X^\dagger = -X\} = \{X \in \mathfrak{g} | Xu = 0\}$ , and  $\mathfrak{g} = \mathfrak{g}_u \oplus \mathfrak{p}$ , where  $\mathfrak{p} = \{X \in \mathfrak{g} | X^\dagger = X\}$ ;*
- (c) *In this case the mapping  $\mathfrak{p} \rightarrow \mathbf{E}$ , given by  $X \mapsto Xu$ , is an isomorphism. Letting  $L_a$  be the unique element of  $\mathfrak{p}$  with  $L_a u = a$ , define*

$$a \bullet b = L_a b$$

*for all  $a, b \in \mathbf{E}$ . Then  $\bullet$  makes  $\mathbf{E}$  a formally real Jordan algebra, with identity element  $u$ .*

<sup>3</sup>Many authors define “reversible” by this condition, i.e. as what we have here called reversible with probability one.

<sup>4</sup>This differs slightly, but not materially, from the definition of an HSD cone in [15], where a fixed inner product is assumed.

*Remark:* The proof of the Koecher-Vinberg Theorem given in [15] takes  $G$  to be the connected identity component of the automorphism group of  $\mathbf{E}$ . We are making the ostensibly stronger claim here that any homogeneously-acting, closed, self-adjoint subgroup of  $\text{Aut}(\mathbf{E})$  will suffice; accordingly, a detailed sketch of the proof is given in an Appendix to this paper.

## 2.3 HSD and Jordan models

We shall say that a model  $A$  is *HSD* (homogeneous and self-dual) iff the cone  $\mathbf{E}_+(A)$  is homogeneous under its group  $\mathcal{G}(A)$  of reversible processes, and equal to its dual with respect to *some* inner product. If  $A$  is an HSD model, then the Koecher-Vinberg theorem implies that  $\mathbf{E}(A)$  carries a unique Euclidean Jordan structure with respect to which the order unit,  $u_A$ , is the identity.

An *idempotent* in  $\mathbf{E}(A)$  is a non-zero element  $p \in \mathbf{E}_+(A)$  such that  $p^2 = p$  (where  $p^2 = p \bullet p$ ). A non-zero idempotent that cannot be decomposed as the sum of two distinct non-zero idempotents is said to be *primitive*. The spectral theorem for Euclidean Jordan algebras (see [15], Proposition III.1.2) tells us that every nonzero element of  $\mathbf{E}_+(A)$  is the sum of positive multiples of pairwise-orthogonal primitive idempotents. It follows that every extremal ray of  $\mathbf{E}(A)_+$  consists precisely of the nonnegative multiples of some primitive idempotent, idempotent generates such an extremal ray. Since the set  $X(A)$  of outcomes of the model  $A$  generates the positive cone  $\mathbf{E}_+(A)$ , we can conclude that every primitive idempotent is a positive multiple of some outcome. However,  $X(A)$  may also contain some non-extremal outcomes. In this section, we identify two simple and natural conditions that together guarantee that every outcome *is*, in fact, a primitive idempotent.

For the balance of this section,  $A$  is an HSD model, equipped with its corresponding Jordan structure and trace, and with the tracial inner product defined by  $\langle a, b \rangle = \text{tr}(ab)$  for all  $a, b \in \mathbf{E}(A)$ . Notice that  $\langle a, b \rangle \geq 0$  for all  $a, b \in \mathbf{E}(A)_+$ . A primitive idempotent  $e \in \mathbf{E}(A)$  satisfies  $\text{tr}(e) = 1$ ; hence, by the Cauchy-Schwarz inequality,  $\langle e, f \rangle \leq 1$  for all primitive idempotents  $f$ . We also have  $\langle e, e \rangle = \langle e, u \rangle = \text{tr}(e) = 1$ . Thus, a primitive idempotent  $e$  defines a pure state,  $\langle e |$  on  $A$ , and this is the unique pure state assigning probability 1 to the effect corresponding to  $e$ .

A *Jordan frame* in a Euclidean Jordan algebra  $\mathbf{E}$  is a set  $e_1, \dots, e_n$  of primitive idempotents summing to  $u$ . All Jordan frames in  $\mathbf{E}$  have the same cardinality, called the *rank* of  $\mathbf{E}$ . By a *Jordan model*, we mean an HSD model such that every outcome is a primitive idempotent, or, equivalently, every test is a Jordan frame.

Let us say that a probabilistic model  $A$  is *uniform* iff there exists a state  $\mu \in \mathbf{E}(A)^*$  taking a constant value  $\mu(x) = 1/m$  on all outcomes  $x \in X(A)$ . Note that this implies that all tests  $E \in \mathfrak{A}(A)$  have cardinality  $m$ . An outcome  $x \in X(A)$  is *unital* iff there exists a state  $\alpha \in \mathbf{E}^*$  with  $\alpha(x) = 1$ , and *sharp* if this state is unique. The model  $A$  itself is unital, respectively, sharp, iff every outcome  $x \in X(A)$  is unital, respectively, sharp. Observe that any Jordan model is sharp (hence, unital) and uniform, with uniform state given by  $\mu(x) = \langle u, x \rangle = 1/n$ ,  $n$  the rank of  $\mathbf{E}$ . We now establish the converse.

**Lemma 5.** *Let  $A$  be HSD.*

- (a) *Every extremal unital outcome is a primitive idempotent.*
- (b) *If  $A$  is uniform, then every unital outcome is extremal, hence, a primitive idempotent.*

*Proof:* (a) Let  $x \in X(A)$  be extremal. As observed above, there exists some  $t > 0$  such that  $tx =: e$ , a primitive idempotent. Now suppose  $f$  is a primitive idempotent representing a pure state of  $\mathbf{E}$ ,

with  $\langle f, x \rangle = 1$ . Then

$$t = t\langle f, x \rangle = \langle f, tx \rangle = \langle f, e \rangle \leq 1,$$

by the Cauchy-Schwarz inequality. Now notice that

$$t^2 \langle x, x \rangle = \langle e, e \rangle = 1$$

so  $\langle x, x \rangle = 1/t^2$ . Choosing any  $E \in \mathfrak{A}(A)$  with  $x \in E$ , we now have

$$\begin{aligned} 1 = \langle e, u \rangle &= t \langle x, u \rangle \\ &= t(\langle x, x \rangle + \sum_{y \in E \setminus \{x\}} \langle x, y \rangle) \\ &\geq t \langle x, x \rangle = t/t^2 = 1/t, \end{aligned}$$

so that  $t \geq 1$ . Thus,  $t = 1$ , and  $x = e$ .

(b) Let  $x \in X(A)$  and  $x = \sum_i s_i x_i$  where the  $x_i$  are extremal outcomes and  $s_i \geq 0$ . Let  $\mu$  be the uniform state on  $\mathbf{E}$ . Then

$$\frac{1}{m} = \mu(x) = \sum_i s_i \mu(x_i) = \sum_i s_i \frac{1}{m}$$

so  $\sum_i s_i = 1$ . If  $x$  is unital, therefore, there exists a pure state assigning probability 1 to  $x$ ; hence, by the self-duality of  $\mathbf{E}(A)_+$ , there exists a primitive idempotent  $f$  with

$$1 = \langle f, x \rangle = \sum_i s_i \langle f, x_i \rangle.$$

Since, as we've just seen,  $s_i \geq 0$  and  $\sum_i s_i = 1$ , we have  $\langle f, x_i \rangle = 1$  for every  $i$  with  $s_i \neq 0$ . But then, every  $x_i$  is a unital extremal outcome and so, by part (a), a primitive idempotent. It follows (again by Cauchy-Schwarz and the argument in the proof of (a)) that  $s_i \neq 0$  implies  $x_i = f$ , whence,  $x = f$ .  $\square$

It follows that any HSD model that is both uniform and unital is a Jordan model. Since, as observed above, the converse also holds, uniform, unital HSD models are exactly the same things as Jordan models. Notice that any Euclidean Jordan algebra  $\mathbf{E}$  can be equipped with the structure of a Jordan model by choosing a distinguished family  $\mathfrak{A}$  of Jordan frames such that the set  $X = \bigcup \mathfrak{A}$  generates  $\mathbf{E}_+$ . In particular, we can always take  $\mathfrak{A}$  to be the set of *all* Jordan frames.

### 3 Composites of Jordan Models

We now wish to examine the structure of composite systems comprising two Jordan models. We begin with a review of the notion of a composite of probabilistic models, following [5, 10].

#### 3.1 Composites and tensor products

Consider two systems  $A$  and  $B$ , which, while possibly interacting, retain enough independence to allow them to be observed and manipulated separately. We would then expect a model for the composite system  $AB$  to include, for each pair of effects  $a \in \mathbf{E}(A)$ ,  $b \in \mathbf{E}(B)$ , a *product effect*  $a \otimes b \in \mathbf{E}(AB)$ , with the understanding that, for a state  $\omega \in \mathbf{E}(AB)^*$ ,  $\omega(a \otimes b)$  gives the joint probability to observe  $a$  and  $b$ . Moreover, we should expect that the two systems can be prepared independently in arbitrary states  $\alpha \in \mathbf{E}(A)^*$ ,  $\beta \in \mathbf{V}(B)$ , so as to produce a *product state*  $\alpha \otimes \beta$



with  $(\alpha \otimes \beta)(a \otimes b) = \alpha(a)\beta(b)$ . Finally, if  $g_A \in G(A)$  and  $g_B \in G(B)$  are symmetries of  $A$  and  $B$ , respectively, then there should exist a symmetry  $g \in G(AB)$  such that  $g(a \otimes b) = ga \otimes gb$  for all  $a \in \mathbf{E}(A)$  and  $b \in \mathbf{E}(B)$ .

Supposing this much, let  $\omega$  be a state on  $\mathbf{E}(AB)$ . We shall say that  $\omega$  is *non-signaling* iff the two *marginal states* given by

$$\omega_A(a) := \sum_{y \in F} \omega(a \otimes y) \quad \text{and} \quad \omega_B(b) := \sum_{x \in E} \omega(x \otimes b)$$

are well-defined, i.e., independent of the choice of tests  $E \in \mathfrak{A}(A)$  and  $F \in \mathfrak{A}(B)$  (This prevents parties controlling  $A$  and  $B$  from sending one another information solely by choosing which tests to measure.) It is not hard to see [36, 9] that this makes the mapping  $a, b \mapsto \omega(a, b)$  bilinear, whence,  $\alpha, \beta \mapsto \alpha \otimes \beta$  and  $a, b \mapsto a \otimes b$  are also bilinear, justifying the tensorial notation. These considerations motivate the following definition.

**Definition 6.** A non-signaling composite of (finite-dimensional) models  $A$  and  $B$  is a model  $AB$ , equipped with two bilinear mappings

$$\otimes : \mathbf{E}(A) \times \mathbf{E}(B) \rightarrow \mathbf{E}(AB) \quad \text{and} \quad \otimes : \mathbf{E}(A)^* \times \mathbf{E}(B)^* \rightarrow \mathbf{E}(AB)$$

such that

- (i) For all tests  $E \in \mathfrak{A}(A)$  and  $F \in \mathfrak{A}(B)$ ,  $E \otimes F = \{x \otimes y | x \in E, y \in F\}$  is a test in  $\mathfrak{A}(AB)$ ;
- (ii)  $(x \otimes y)(\alpha \otimes \beta) = \alpha(x)\beta(y)$  for all states  $\alpha \in \mathbf{E}(A)^*, \beta \in \mathbf{E}(B)^*$ ;
- (iii) For all  $\tau_A \in \mathcal{D}(A)$  and  $\tau_B \in \mathcal{D}(B)$ , there exists a process  $\tau \in \mathcal{G}(AB)$  such that

$$\tau(\alpha \otimes \beta) = \tau_A \alpha \otimes \tau_B \beta$$

for all  $\alpha \in \mathbf{E}(A)^*, \beta \in \mathbf{E}(B)^*$ .

It follows from (i) that if  $x$  and  $y$  are outcomes of  $A$  and  $B$ , then  $x \otimes y$  is an outcome of  $AB$ . This, together with condition (ii) and the bilinearity of the mappings  $\otimes$ , that  $(\alpha \otimes \beta)(a \otimes b) = \alpha(a)\beta(b)$  for all  $a \in \mathbf{E}(A), b \in \mathbf{E}(B)$  and all  $\alpha \in \mathbf{E}(A)^*, \beta \in \mathbf{E}(B)^*$ . We also have  $u_A \otimes u_B = u_{AB}$ .

**Definition 7.** We say that  $AB$  is locally tomographic iff every bipartite state  $\omega \in \mathbf{E}(AB)_+^*$  is entirely determined by the joint probabilities  $\omega(a, b) := \omega(a \otimes b)$  that  $\omega$  assigns to pairs of effects  $a \in \mathbf{E}(A), b \in \mathbf{E}(B)$ .

If  $A$  and  $B$  are finite-dimensional (that is, if  $\mathbf{E}(A)$  and  $\mathbf{E}(B)$  are finite-dimensional), the condition that a non-signaling composite be local tomographic is equivalent to the condition that  $\dim(\mathbf{E}(AB)^*) = \dim(\mathbf{E}(A)^*) \dim(\mathbf{E}(B)^*)$ , that is, as vector spaces (ignoring the order structure)  $\mathbf{E}(AB)^* = \mathbf{E}(B)^* \otimes \mathbf{E}(A)^*$  and  $\mathbf{E}(AB) = \mathbf{E}(A) \otimes \mathbf{E}(B)$ . Note that this makes the process  $\tau \in \mathcal{D}(AB)$  required by condition (iii) above unique, so that we can sensibly write  $\tau = \tau_A \otimes \tau_B$ .

### 3.2 Proof of Theorem 1

We now consider the implications of the existence of a locally tomographic HSD composite,  $AB$ , of HSD systems  $A$  and  $B$ . Recalling the notation used in Section 3, if  $A$  is any system, let  $G(A)$  denote the connected identity component of the group  $\mathcal{G}(A)$  of invertible physical processes on  $A$ . If  $\mathcal{G}(A)$  acts homogeneously on  $\mathbf{E}_+$ , so does  $G(A)$  ([15], p. 5).

Since  $AB$  is HSD, we can introduce an inner product  $\langle, \rangle_{AB}$  on  $\mathbf{E}(AB)$  that is normalized and self-dualizing for  $\mathbf{E}(AB)_+$ . We shall say that  $\langle, \rangle_{AB}$  *factors*, and that  $\mathbf{E}(AB)$  is *factorizably self-dual*, iff

$$\langle a \otimes b, c \otimes d \rangle = \langle a, c \rangle \langle b, d \rangle$$

for all  $a, c \in \mathbf{E}(A)$  and  $b, d \in \mathbf{E}(B)$  where  $\langle, \rangle_A$  and  $\langle, \rangle_B$  are self-dualizing inner products on  $\mathbf{E}(A)$  and  $\mathbf{E}(B)$ , respectively. More generally, say that  $AB$  is *factorizably self-dual* iff the self-dualizing inner product can be chosen to factor in this way. This is not as restrictive a condition as it might at first seem:

**Lemma 8.** *Let  $AB$  be a non-signaling composite of Jordan models  $A$  and  $B$ . If  $AB$  is itself Jordan, then the trace form on  $\mathbf{E}(AB)$  factors.*

*Proof:* By the definition of a composite, if  $x, y \in X(A)$ , then  $x \otimes y$  is an outcome in  $X(AB)$ . Since  $x$  and  $y$  are unital in  $A$  and  $B$ ,  $x \otimes y$  is unital in  $X(AB)$ : the product state  $\langle x | \otimes \langle y |$  assigns  $x \otimes y$  probability 1 (again, by the definition of a composite). Hence, by Lemma 6 (b),  $x \otimes y$  is a primitive idempotent in  $\mathbf{E}(AB)$ , and therefore pure. But since  $AB_+$  is HSD, there is a unique pure state,  $\langle x \otimes y |$ , with  $\langle x \otimes y | x \otimes y \rangle = 1$ . Hence,  $\langle x | \otimes \langle y | = \langle x \otimes y |$ , so that  $\langle x \otimes y | a \otimes b \rangle = \langle x | a \rangle \langle y | b \rangle$  for all  $a \in \mathbf{E}(A), b \in \mathbf{E}(B)$ . Since  $X(A)$  spans  $\mathbf{E}(A)$  and  $X(B)$  spans  $\mathbf{E}(B)$  the same holds with arbitrary elements of  $\mathbf{E}(A)$  and  $\mathbf{E}(B)$  in place of  $x$  and  $y$  respectively, i.e, the inner product factors.  $\square$

*Remark:* Rather than assuming that  $AB$  is Jordan (equivalently, HSD, uniform, and unital), one can suppose that it is HSD and that every  $E \in \mathfrak{A}(AB)$  has the same cardinality. Since  $\mathfrak{A}(AB)$  contains product tests, the cardinality of its tests must then be  $mn$  where  $m$  and  $n$  are the ranks of the Jordan algebras  $\mathbf{E}(A)$  and  $\mathbf{E}(B)$ . Now the product of the uniform state on  $A$  with the uniform state on  $B$  provides a uniform state on  $AB$ . Lemma 6 can then be invoked, as in the proof above.

**Lemma 9.** *Let  $AB$  be locally tomographic and factorizably self-dual.*

- (a) *If  $g \in GL(\mathbf{E})$ , then  $(g \otimes \mathbf{1}_B)^\dagger = g^\dagger \otimes \mathbf{1}_B$*
- (b) *If  $g \in G(A)$ , then  $g^\dagger \otimes \mathbf{1}_B$  is an order-automorphism of  $\mathbf{E}(AB)$ .*

*Proof:* (a) For elements of  $\mathbf{E}(AB)$  of the form  $a \otimes b$ , we have

$$\langle (g \otimes \mathbf{1}_B) a_1 \otimes b_1, a_2 \otimes b_2 \rangle = \langle g a_1, a_2 \rangle \langle b_1, b_2 \rangle = \langle a_1, g^\dagger a_2 \rangle \langle b_1, b_2 \rangle = \langle a_1 \otimes b_1, (g^\dagger \otimes \mathbf{1}_B) a_2 \otimes b_2 \rangle.$$

Since  $AB$  is locally tomographic, such elements span  $\mathbf{E}(AB)$ , so the relation holds generally.

(b) This now follows, since the adjoint of an order-automorphism with respect to a self-dualizing inner product is again an order-automorphism (cf. [15], I.1.7).  $\square$

Let  $AB$  be a locally tomographic, HSD composite of HSD models  $A$  and  $B$ , and suppose  $\langle, \rangle_{AB}$  is a factorizable self-dualizing inner product on  $\mathbf{E}(AB)$ , where the inner products on the factors are chosen in accordance with Theorem 8 (a). Let  $G(A)^\dagger$  denote the group of adjoints  $g^\dagger$  where  $g \in G(A)$ , with the adjoint defined by the chosen inner product. As noted above, the self-duality of  $\mathbf{E}(A)$  ensures that  $G(A)^\dagger \leq \text{Aut}(\mathbf{E})$ . Now let

$$G_A := \langle G(A) \cup G(A)^\dagger \rangle,$$

that is,  $G_A$  is the closed subgroup of  $\text{Aut}(\mathbf{E})$  generated by  $G(A)$  and  $G(A)^\dagger$ . The group  $G_A$  acts transitively on the interior of  $\mathbf{E}_+$ , since  $G(A)$  does, and is easily seen to be self-adjoint and connected. Theorem 8 therefore yields, for each  $a \in \mathbf{E}(A)$ , a unique self-adjoint element  $L_a$  of the Lie algebra  $\mathfrak{g}_A$  of  $G_A$ , such that  $L_a u_A = a$ , and such that  $a \bullet b := L_a b$  defines the Jordan structure of  $\mathbf{E}(A)$ . Moreover, we have

**Lemma 10.** *If  $g \in G_A$ , then  $g \otimes \mathbf{1}_B \in G_{AB}$ .*

*Proof:* For every element  $g \in G(A)$ ,  $g \otimes \mathbf{1}_B \in G(AB)$  (by condition (iii) of Definition 9), whence, by Lemma 10,  $g^\dagger \otimes \mathbf{1}_B$  belong to  $G(AB)^\dagger$ .  $\square$

Thus, we have a canonical embedding  $G_A \simeq G_A \otimes \{\mathbf{1}_B\} \leq G_{AB}$ . As in Theorem 4, let  $\mathfrak{g}_A$  denotes the Lie algebra of  $G_A$  and  $\mathfrak{p}_A$ , the self-adjoint part of  $G_A$ , and similarly for  $\mathfrak{g}_{AB}$  and  $\mathfrak{p}_{AB}$ .

**Lemma 11.** *For all  $a \in \mathbf{E}(A), v \in \mathbf{E}(B)$ ,  $L_a \otimes \mathbf{1}_B, \mathbf{1}_A \otimes L_v \in \mathfrak{g}_{AB}$ . Hence,  $L_{a \otimes u_B} = L_a \otimes \mathbf{1}$  and  $L_{u_A \otimes v} = \mathbf{1}_A \otimes L_v$ .*

*Proof:* Since  $G_A \leq G_{AB}$  (via  $g \mapsto g \otimes \mathbf{1}_B$ ), and as  $L_a \in \mathfrak{p}_A$ , we have a one-parameter group  $g : t \mapsto g_t \in G_A$  with  $L_a = g'(0)$ , and a corresponding one-parameter group  $g \otimes \mathbf{1} : t \mapsto g_t \otimes \mathbf{1}$  in  $G_{AB}$ . The bilinearity of the tensor product gives us  $L_a \otimes \mathbf{1} = (g \otimes \mathbf{1})'(0) \in \mathfrak{g}_{AB}$ . Since the inner product factors, we have

$$(L_a \otimes \mathbf{1})^\dagger = (L_a)^\dagger \otimes \mathbf{1}^\dagger = L_a \otimes \mathbf{1}$$

So  $L_a \otimes \mathbf{1}$  is a self-adjoint element of  $\mathfrak{g}_{AB}$ , that is, an element of  $\mathfrak{p}_{AB}$ . Also,  $(L_a \otimes \mathbf{1})(u_A \otimes u_B) = L_a u_A \otimes u_B = a \otimes u_B$ . The second identity is proved similarly.  $\square$

**Corollary 12.** *In  $AB$ , we have*

$$(a \otimes u_B) \bullet (b \otimes v) = (a \bullet b) \otimes v \quad \text{and} \quad (u_A \otimes v) \bullet (a \otimes w) = a \otimes (v \bullet w).$$

*Proof:* Taking the first identity, we have

$$(a \otimes u_A) \bullet (b \otimes v) = L_{a \otimes u_B}(b \otimes v) = (L_a \otimes \mathbf{1}_B)(b \otimes v) = L_a b \otimes v = (a \bullet b) \otimes v.$$

Similarly for the second identity.  $\square$

This corollary is what we promised to establish, namely that a locally tomographic, factorizably HSD composite of HSD models must satisfy Hanche-Olsen's condition, (1) — hence, by Hanche-Olsen's *theorem*, the models involved are self-adjoint parts of complex  $C^*$ -algebras, giving Proposition 1.<sup>5</sup> In the language of Section 2.2, and appealing to Lemma 10, we can rephrase this as asserting that *any locally tomographic, non-signaling theory in which all models are Jordan models, is a standard quantum theory*.

In the next section, we consider the implications of this result in the setting of a dagger-monoidal category  $\mathcal{C}$  of HSD probabilistic models.

## 4 Categorical Considerations

It is reasonable to represent a physical theory as a category,  $\mathcal{C}$ , in which objects represent distinct physical systems and morphisms represent physical processes. To capture the idea that processes can be composed not only serially, but also in parallel, it's equally natural to suppose that  $\mathcal{C}$  carries a symmetric monoidal structure. This point of view has been developed extensively by Abramsky

---

<sup>5</sup>The reader may have noticed that the only explicit use we've made of local tomography in the derivation of Corollary 12 is in the proof of part (a) of Lemma 11. However, to apply Hanche-Olsen's theorem together with Corollary 12, we need to know that the ordinary vector space tensor product  $\mathbf{E}(A) \otimes \mathbf{E}(B)$ , where  $B$  is a qubit, has an HSD structure; this only follows if the composite  $AB$  is locally tomographic.

and Coecke [1], Baez [4], and Selinger [33]. A striking result of this work is that many qualitative features of quantum information processing are actually direct consequences of the fact that the category of finite-dimensional Hilbert spaces and linear mappings is a *dagger-compact* category.

We recall that a *symmetric monoidal category* is a category  $\mathcal{C}$ , equipped with a bi-functorial operation  $\otimes : \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$  that is commutative and associative, up to natural isomorphisms  $A \otimes B \simeq B \otimes A$  and  $A \otimes (B \otimes C) \simeq (A \otimes B) \otimes C$  for objects  $A, B, C$  of  $\mathcal{C}$ , and also equipped with a *tensor unit*,  $I$ , such that  $I \otimes A \simeq A$  for all  $A \in \mathcal{C}$ . We shall be interested here in symmetric monoidal categories having probabilistic models as objects. It would be natural to take morphisms to be those positive linear mappings that we wish to consider physical processes; for computational convenience, however, it seems reasonable to enlarge the set of morphisms to include arbitrary linear combinations of such processes. This suggests the following:

**Definition 13.** A finite-dimensional *monoidal probabilistic theory* is a symmetric monoidal category  $\mathcal{C}$  in which

- (i) Objects are finite-dimensional probabilistic models;
- (ii) For all objects  $A, B \in \mathcal{C}$ , the set  $\mathcal{C}(A, B)$  of morphisms  $A \rightarrow B$  is a non-trivial subspace of the space  $\mathcal{L}(\mathbf{E}(A), \mathbf{E}(B))$  of linear mappings  $\phi : \mathbf{E}(A) \rightarrow \mathbf{E}(B)$ , equipped with a generating cone  $\mathcal{C}_+(A, B)$  of positive linear mappings. Composition of morphisms is composition of mappings, and  $\mathcal{C}_+(A, B) \circ \mathcal{C}_+(B, C) \subseteq \mathcal{C}_+(A, C)$  for all  $A, B, C \in \mathcal{C}$ ;
- (iii) The tensor unit  $I$  is  $\mathbb{R}$  (with its usual order and unit), and  $\mathcal{C}(I, A) = \mathbf{E}(A)$ .
- (iv) For every  $A \in \mathcal{C}$ ,  $\mathcal{D}(A)$  is the set of all morphisms  $\phi \in \mathcal{C}_+(A, A)$  satisfying  $\phi(u_A) \leq u_B$ . Accordingly, the group  $G(A)$  generated by reversible processes in  $\mathcal{D}(A)$ , is exactly the group of invertible elements of  $\mathcal{C}(A, A)$  having inverses in  $\mathcal{C}_+(A, A)$ .
- (v) The monoidal product,  $AB$ , of two objects  $A, B \in \mathcal{C}$ , is a locally tomographic composite of the models  $A$  and  $B$ , in the sense of Definition 8, and the monoidal operation  $\mathcal{C}(A, A') \times \mathcal{C}(B, B') \rightarrow \mathcal{C}(AA', BB')$  is bilinear for all  $A', B' \in \mathcal{C}$ .

By condition (ii),  $\mathcal{C}(A, I)$  is a non-trivial subspace of  $\mathbf{E}(A)^*$ , but, absent some further constraint, these need not be isomorphic. Indeed, the requirement that  $\mathcal{C}(A, I) \simeq \mathbf{E}(A)^*$  is equivalent to requiring that every state  $\alpha$  on  $A \in \mathcal{C}$  correspond to a morphism  $\alpha : A \rightarrow I$ , hence, to an element of  $\mathbf{E}(A)^*$ . This is precisely the state-completeness assumption discussed in Section 2. Given state-completeness, if  $\alpha \in \mathbf{E}(A)^* \simeq \mathcal{C}(A, I)$  and  $a \in \mathbf{E}(A) \simeq \mathcal{C}(I, A)$ , we have

$$\alpha(a) = \alpha \circ a.$$

**Lemma 14.** Let  $\mathcal{C}$  be a state-complete monoidal probabilistic theory. Then for every pair of objects  $A, B \in \mathcal{C}$ , the composite  $AB$  is non-signaling.

*Proof:* If  $\mathcal{C}$  is state-complete, the linearity of morphisms and the bilinearity of  $\otimes$  together imply that composite systems in  $\mathcal{C}$  are non-signaling. Suppose  $\omega : A \rightarrow I$ . If  $E \in \mathfrak{A}(A)$  and  $y \in X(B)$ , then — identifying each  $x \in E$  with the corresponding linear mapping  $x : I \rightarrow A$ , and similarly for  $y$  — we have

$$\sum_{x \in E} \omega(x, y) := \sum_{x \in E} \omega \circ (x \otimes y) = \omega\left(\sum_{x \in E} x \otimes y\right) = \omega(u_A, y)$$

Similarly,  $\sum_{y \in F} \omega(x, y) = \omega(x, u_B)$  for all  $F \in \mathfrak{A}(B)$ .  $\square$

From this point on, we assume that  $\mathcal{C}$  is state-complete.

A *dagger* [33] on a category  $\mathcal{C}$  is a contravariant endo-functor<sup>6</sup>  $\dagger : \mathcal{C} \rightarrow \mathcal{C}$  such that, for all objects  $A \in \mathcal{C}$ ,  $A^\dagger = A$ , and, for all morphisms  $\phi, \psi$  in  $\mathcal{C}$ ,  $\phi^{\dagger\dagger} = \phi$ . An morphism  $\phi$  in  $\mathcal{C}$  is *unitary* with respect to  $\dagger$  iff it is invertible and satisfies  $\phi^{-1} = \phi^\dagger$ . A *dagger-monoidal category* is a symmetric monoidal category equipped with a dagger that commutes with the monoidal structure, so that  $(\phi \otimes \psi)^\dagger = \phi^\dagger \otimes \psi^\dagger$  for all morphisms  $\phi$  and  $\psi$  in  $\mathcal{C}$ , and is such that the isomorphisms comprising the components of the natural associativity, symmetry, and unit-introduction transformations are unitary.

Let  $\mathcal{C}$  be a monoidal probabilistic theory, equipped with a dagger. Assume, further, that the dagger operation is linear, and positive; that is, if  $\phi \in \mathcal{C}_+(A, B)$ , then  $\phi^\dagger \in \mathcal{C}_+(B, A)$ . In this setting, we have  $\mathcal{C}(A, I) \simeq \mathcal{C}(I, A) = \mathbf{E}(A)$ , whence, by dimensional considerations,  $\mathcal{C}(A, I) \simeq \mathbf{E}(A)^*$  as a linear space.<sup>7</sup> also have  $\mathcal{C}_+(A, I) \simeq \mathcal{C}(I, A)$ , but the possibility remains open that  $\mathcal{C}_+(A, I)$  is a proper sub-cone of the dual cone  $\mathbf{E}(A)_+^*$ .

The dagger also provides us with a canonical bilinear form on each  $\mathbf{E}(A)$ ,  $A \in \mathcal{C}$ , defined by

$$\langle a, b \rangle_A := b^\dagger \circ a$$

for all  $a, b \in \mathcal{C}(I, A) \simeq \mathbf{E}(A)$ . Since  $r^\dagger = r$  for every  $r \in \mathbb{R} = I$ , this form is symmetric:

$$\langle b, a \rangle = a^\dagger \circ b = (b^\dagger \circ a)^\dagger = \langle a, b \rangle^\dagger = \langle a, b \rangle.$$

An element  $g$  of the group  $\mathcal{G}(A)$  of invertible morphisms in  $\mathcal{C}(A, A)$  is said to be *unitary* iff  $g^\dagger = g^{-1}$ . We write  $U(A)$  for the group of unitary elements of  $\mathcal{G}(A)$ . The bilinear form  $\langle \cdot, \cdot \rangle_A$  is invariant with respect to  $U(A)$ , in the sense that

$$\langle ga, gb \rangle_A = (g \circ b)^\dagger \circ g \circ a = b^\dagger \circ g^\dagger \circ g \circ a = b^\dagger \circ a = \langle a, b \rangle_A.$$

This bilinear form is also *positive*, in the sense that  $\langle a, b \rangle \geq 0$  for  $a, b \in A_+$ .<sup>8</sup> Finally, observe (from the bifunctoriality of  $\otimes$ ) that the canonical bilinear form *factors*, in the sense that, for every  $A, B \in \mathcal{C}$ ,

$$\langle a \otimes b, c \otimes d \rangle_{AB} = \langle a, c \rangle_A \langle b, d \rangle_B \quad (2)$$

for all  $a, c \in \mathcal{C}(I, A) \simeq \mathbf{E}(A)$  and all  $b, d \in \mathcal{C}(I, B) \simeq \mathbf{E}(B)$ .

**Definition 15.** A *dagger-monoidal probabilistic theory* is a (state-complete) monoidal probabilistic theory  $\mathcal{C}$  equipped with a (positive, linear) dagger, such that, for every  $A \in \mathcal{C}$ , and every  $g \in U(A)$ ,  $gu_A = u_A$ . We shall call  $\mathcal{C}$  *dagger-HSD* iff (i) the canonical form  $\langle \cdot, \cdot \rangle_A$  is an inner product (i.e., is positive semidefinite) and (ii) every system  $A$  in  $\mathcal{C}$  is homogeneous with respect to  $\mathcal{C}(A, A)$ .

*Remark:* The condition that  $\langle \cdot, \cdot \rangle$  be an inner product is non-trivial, but can be motivated in the case in which every system  $A \in \mathcal{C}$  is an irreducible Jordan model. In this case, using [38], Corollary 2 and Lemma 6, one can show that any positive,  $G(AA)$ -invariant bilinear form on  $AA$  will in fact be an inner product.

Let  $\mathcal{C}$  be any dagger-HSD probabilistic theory. It is straightforward to verify that, if  $A \in \mathcal{C}$  and  $\tau \in \mathcal{C}(A, A)$ , then  $\tau^\dagger \in \mathcal{C}(A, A)$  functions as the adjoint of  $a$  with respect to the canonical inner product. That is, if  $a, b \in \mathcal{C}(I, A) \simeq A$ , then

$$\langle \tau a, b \rangle = (\tau \circ a)^\dagger \circ b = a \circ \tau^\dagger \circ b = \langle a, \tau^\dagger b \rangle.$$

<sup>6</sup>A contravariant endo-functor  $F$  on a category  $\mathcal{C}$  is a map from the morphisms of  $\mathcal{C}$  to the morphisms of  $\mathcal{C}$ , such that  $F(f \circ g) = F(g) \circ F(f)$  (as opposed to an ordinary (covariant) endo-functor, for which  $F(f \circ g) = F(f) \circ F(g)$ ).

<sup>7</sup>Note that we take the states to be elements of  $\mathcal{C}(A, I)$ , while the effects belong to  $\mathcal{C}(I, A)$ . This is the reverse of the convention in many papers (e.g. [1, 33], and ourselves in [8]), and may be thought of as working in the “generalized Heisenberg picture” for probabilistic theories. It is motivated in part by the naturality of viewing states as functionals from an order-unit space (or test space) to  $\mathbb{R}$ , and effects as dual to these.

<sup>8</sup>This is quite distinct from positive-definiteness, i.e., we are not claiming that  $\langle a, a \rangle > 0$  for all  $a \neq 0$ .

In particular, then, the group  $\mathcal{G}(A)$  is self-adjoint with respect to  $\langle \cdot, \cdot \rangle_A$  (i.e.  $x \in G(A) \implies x^\dagger \in G(A)$ ). It follows that the cone  $\mathbf{E}(A)_+$  is self-dual (cf. [15], Exercise I.8).

In view of Proposition 1, the factorization property (2) immediately yields the corollary that every locally tomographic dagger-HSD theory containing at least one qubit, is a standard quantum theory in the sense that all of its systems are isomorphic, as ordered linear spaces, to the self-adjoint parts of complex matrix algebras with their standard orderings, and the monoidal product is the usual tensor product of quantum systems. That is,

**Proposition 16.** *A locally tomographic, dagger-HSD probabilistic theory in which at least one system has the structure of a qubit, is a standard quantum theory.*

## 5 Conclusions

We have shown that, in the specific context of (finite-dimensional, state-complete) Jordan probabilistic models — that is, uniform, unital models with homogeneous, self-dual cones — defining features of orthodox, complex QM are

- (i) the availability of locally tomographic, non-signaling products — otherwise, a weak constraint;
- (ii) the existence of a qubit.

Similarly, in the context of a state-complete dagger-HSD theory (where composites are automatically non-signaling), quantum theory is picked out by local tomography and the existence of a qubit.

In [12], Dakić and Brukner derive QM from assumptions that also include (in effect) the existence of a qubit; however, they make a very strong uniformity assumption, namely, that all systems of a given information-carrying capacity, are isomorphic. This is not unreasonable if we imagine that all systems are built up, through a uniform process of composition, from a single elementary system — in this case, a qubit. And, indeed, in [29], it is shown that any probabilistic theory of the general type considered here, in which every system arises as a non-signaling, locally tomographic product of qubits, and in which, for every system  $A$ , the group  $G(A)_u$  acts continuously on the set of pure states, is quantum. In contrast, our approach shows that, in the context of a probabilistic theory in which systems are represented by Jordan models, the mere *existence* of a *single* qubit, together with the possibility of forming locally tomographic, non-signaling composite systems, is enough to enforce all the structure of QM, including the aforementioned uniformity assumption. We have a similar result for any dagger-HSD theory.

Even so, various interesting questions remain regarding HSD theories. For one thing, it would be very interesting to understand the possibilities for *non*-locally tomographic composites in such a theory: this should shed light on real and quaternionic QM, in particular. In a different direction, one would want to know whether it is possible to weaken, or entirely to dispense with, the assumption in Theorem 1, that the theory  $\mathcal{C}$  includes a qubit. If so, then in the context of locally tomographic, non-signaling probabilistic theories in which systems are unital and uniform, *or* the context of dagger-monoidal probabilistic theories in which the canonical bilinear form is positive-definite, the HSD condition by itself is sufficient to rule out non- $C^*$ -algebraic theories. Of course, it would be at least equally interesting to construct a non- $C^*$ -algebraic, locally-tomographic, non-signaling HSD theory that *not* contain a qubit.

*Acknowledgement:* We thank C. M. Edwards for drawing our attention to Hanche-Olsen’s paper. Part of this work was done while the authors were guests of the Oxford University Computing Laboratory, whose hospitality is also gratefully acknowledged. H. B. thanks the Foundational Questions

Institute (FQXi) for travel support for the visit. Additional work was done at the Perimeter Institute for Theoretical Physics; work at Perimeter Institute is supported in part by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation.

## References

- [1] S. Abramsky and B. Coecke. A categorical semantics of quantum protocols. *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science (LICS '04)*, pages 415–425, 2004.
- [2] E. Alfsen and F. W. Shultz, *Geometry of state spaces of operator algebras*. Birkhäuser (2003).
- [3] H. Araki, On a characterization of the state space of quantum mechanics, *Comm. Math. Phys.* **75** (1980), 1-24.
- [4] J. Baez, Quantum Quandaries: a category-theoretic perspective, in D. Rickles, S. French and J. Saatsi, *The Structural Foundations of Quantum Gravity*, Oxford, 2006 (arXiv: [arxiv.org/abs/quant-ph/0404040v2](https://arxiv.org/abs/quant-ph/0404040v2), 2004)
- [5] H. Barnum, J. Barrett, M. Leifer and A. Wilce, A generalized no-broadcasting theorem, *Phys. Rev. Lett.* **99**, 240501-240504 (2007)
- [6] H. Barnum, J. Barrett, M. Leifer and A. Wilce, Teleportation in general probabilistic theories, in Proceedings of the Clifford Lectures, Tulane University, March 12-15, 2008, to appear in *Proceedings of Symposia in Applied Mathematics* (AMS); also [arXiv:0805.3553](https://arxiv.org/abs/0805.3553) (2008).
- [7] H. Barnum, P. Gaebler and A. Wilce, Ensemble steering, weak self-duality, and the structure of probabilistic theories, [arXiv:0912.5532](https://arxiv.org/abs/0912.5532) (2009)
- [8] H. Barnum, R. Duncan, A. Wilce, Symmetry, compact closure, and dagger compactness for categories of convex operational models, e-print [arxiv:1004.2920](https://arxiv.org/abs/1004.2920) (2010). Presented at QPL VII, Oxford, May 29-30, 2010.
- [9] H. Barnum, C. Fuchs, J. Renes and A. Wilce, Influence-free states on compound quantum systems, [arXiv:quant-ph/0507108](https://arxiv.org/abs/quant-ph/0507108). (2005).
- [10] H. Barnum and A. Wilce, Ordered linear spaces and categories as frameworks for information-processing characterizations of quantum theory, [arxiv:0908.2354](https://arxiv.org/abs/0908.2354), 2009.
- [11] J. Barrett, Information processing in generalized probabilistic theories, *Physical Review A* **75**, 032304(2007) ([arXiv:quant-ph/0508211v3](https://arxiv.org/abs/quant-ph/0508211v3), 2005)
- [12] B. Dakić and Č. Brukner, Quantum theory and beyond: is entanglement special? [arXiv:0911.0695](https://arxiv.org/abs/0911.0695) (2009)
- [13] E. B. Davies and J. T. Lewis An operational approach to quantum probability, *Comm. Math. Phys.* **17** (1970), 239-260.
- [14] G. Chiribella, G. M. D’Ariano and P. Perinotti, Reversible realization of physical processes in probabilistic theories (2009) ([arXiv:0908.1583](https://arxiv.org/abs/0908.1583)). Published version: Probabilistic theories with purification, *Phys. Rev. A* **81**, 062348 (2010).
- [15] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford, University Press (1994).
- [16] D. Foulis and C. Randall, Empirical logic and tensor products, in H. Neumann (ed.), *Interpretations and Foundations of Quantum Theory*, B. I. Wissenschaft, Mannheim (1981).

- [17] P. Goyal, From information geometry to quantum theory, *New J. Phys.* **12** (2010), 023012.
- [18] H. Hanche-Olsen, JB-algebras with tensor products are  $C^*$ -algebras, in H. Araki et al., (eds.), *Operator Algebras and their Connections with Topology and Ergodic Theory*, Lecture Notes in Mathematics 1132, Springer, 1985.
- [19] L. Hardy, Quantum theory from five reasonable axioms, [quant-ph/00101012](#) (2000).
- [20] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Mechanics*, North-Holland, 1982. (2nd edition: Edizioni della Normale, Pisa, 2011).
- [21] P. Jordan, Ueber verallgemeinerungsmöglichkeiten des formalismus der quantenmechanik, *Nachr. Akad. Wiss. Göttingen Math. Phys. Kl.*, **41** (1933) 209-217.
- [22] P. Jordan, J. von Neumann and E. P. Wigner, On an algebraic generalization of the quantum-mechanical formalism, *Annals of Mathematics* **35** (1934) 29-64.
- [23] A. Knapp, *Lie Groups Beyond an Introduction*, 2nd ed., Birkhauser, 2002
- [24] M. Koecher, Die geodätischen von positivitätsbereichen, *Math. Annalen* **135** (1958) 192-202.
- [25] G. Ludwig, *Foundations of Quantum Mechanics*, Springer, 1985
- [26] G. Mackey, *Mathematical Foundations of Quantum Mechanics*, Addison Wesley, 1963
- [27] Ll. Masanes and M. Müller, A derivation of quantum theory from physical requirements, *New J. Phys.*, **13** (2011) ([arXiv:1004.1483](#), 2011)
- [28] M. Müller and C. Ududec, The computational power of quantum mechanics determines its self-duality, [arXiv:1110.3516](#) (2011)
- [29] G. de la Torre, Ll. Masanes, A. Short and M. Müller, Deriving quantum theory from its local structure and reversibility, [arXiv:1110.5482](#) (2011)
- [30] J. Rau, On quantum vs. classical probability, *Annals of Physics* **324** (2009) 2622–2637
- [31] I. Satake, *Algebraic structures of symmetric domains*, (Publications of the Mathematical Society of Japan, no. 14), Iwanami Shoten and Princeton University Press, 1980.
- [32] E. Schrödinger, Probability relations between separated systems, *Proceedings of the Cambridge Philosophical Society* **32** 446-452 (1936).
- [33] P. Selinger. Towards a semantics for higher-order quantum computation. In *Proceedings of the 2nd International Workshop on Quantum Programming Languages*, Turku Finland, pages 127–143. Turku Center for Computer Science, 2004. Publication No. 33.
- [34] E. B. Vinberg, Homogeneous cones, *Dokl. Acad. Nauk. SSSR* **141** (1960) 270-273; English trans. *Soviet Math. Dokl.* **2** (1961) 1416-1619.
- [35] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton, 1955
- [36] A. Wilce, The tensor product in generalized measure theory, *Int. J. Theor. Phys.* Volume 31, Number 11, 1915-1928 (1992)
- [37] A. Wilce, Four and a half axioms for finite-dimensional quantum theory, in Y. Ben-Menahem and M. Hemmo (eds.) *Probability in Physics: essays in honor of Itamar Pitowsky*, 2012 ([arXiv:0912.5530](#), 2009).
- [38] A. Wilce, Symmetry, self-duality, and the Jordan structure of quantum theory, [arXiv:1110.6607](#) (2011)



## Appendix: The Koecher-Vinberg Theorem

This appendix contains a detailed sketch of the proof of the Koecher-Vinberg Theorem. This is almost entirely extracted from Faraut and Koranyi [15], to whom we refer for many of the details, but with a few minor modifications in order to obtain the precise form of the theorem (our Theorem 8) that we required above.

In what follows,  $\mathbf{E}$  is a finite-dimensional order-unit space with a self-dual positive cone  $\mathbf{E}_+$ . By this we mean that *there exists* a self-dualizing inner product on  $\mathbf{E}$ . Let  $\text{Aut}(\mathbf{E})$  denote the group of order-automorphisms of  $\mathbf{E}$ . This is a Lie group, as is any closed subgroup  $G \leq \text{Aut}(\mathbf{E})$ .

When  $G$  is connected and acts homogeneously on  $\mathbf{E}_+$  (that is, transitively on the interior of  $\mathbf{E}_+$ ), we can use this action to construct a Jordan product on  $\mathbf{E}$ , as per Theorem 8, which, for convenience, we now restate:

**Theorem:** *Let  $G$  be a closed, connected subgroup of  $\text{Aut}(\mathbf{E})$  and let  $\mathfrak{g}_u$  denote the Lie algebra of  $G_u$ , the stabilizer of  $u$  in  $G$ . Then*

- (a) *It is possible to choose a self-dualizing inner product on  $\mathbf{E}_+$  in such a way that  $G_u = G \cap O(\mathbf{E})$ , where  $O(\mathbf{E})$  is the orthogonal group with respect to the chosen inner product;*
- (b) *If  $G = G^\dagger$  with respect to this inner product, then*

$$\mathfrak{g}_u = \{X \in \mathfrak{g} \mid X^\dagger = -X\} = \{X \in \mathfrak{g} \mid Xu = 0\},$$

*and  $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{g}_u$ , where  $\mathfrak{p} = \{X \in \mathfrak{g} \mid X^\dagger = X\}$ .*

- (c) *In this case the mapping  $\mathfrak{p} \rightarrow \mathbf{E}$ , given by  $X \mapsto Xu$ , is an isomorphism of linear spaces. Letting  $L_a$  denote the unique element  $X \in \mathfrak{p}$  with  $Xu = a$ , define*

$$a \bullet b = L_a b$$

*for all  $a, b \in \mathbf{E}$ . Then  $\bullet$  makes  $\mathbf{E}$  a formally real Jordan algebra, with identity element  $u$ .*

We break the proof into a series of Lemmas. Throughout,  $G$  is a connected, closed subgroup of  $\text{Aut}(\mathbf{E})$ , acting homogeneously on the interior,  $\mathbf{E}_+^\circ$ , of the cone  $\mathbf{E}_+$ .

**Lemma A:** *If  $g \in \text{Aut}(\mathbf{E})$ , then  $g^* \in \text{Aut}(\mathbf{E})$ , where  $g^*$  is the adjoint with respect to any self-dualizing inner product.*

*Proof:* If  $g \in \text{Aut}(\mathbf{E})$  preserves  $\mathbf{E}_+$ , then  $g^\dagger$  preserves  $\mathbf{E}_+^\perp = \mathbf{E}_+$ .  $\square$

**Lemma B:** *Any compact subgroup of  $\text{Aut}(\mathbf{E})$  fixes some point  $a$  in the interior of  $\mathbf{E}_+$ . In particular, a maximal compact subgroup is a stabilizer, and vice versa. Thus, all maximal compact subgroups of  $\text{Aut}(\mathbf{E})$  are conjugate.*

For a proof, see [15], Proposition I.1.8ff.

**Lemma C:** *For a suitable choice of self-dualizing inner product,  $O(\mathbf{E}) \cap G \leq G_u$ , where  $O(\mathbf{E})$  is the orthogonal group relative to the chosen inner product.*

*Proof:* If  $\langle \cdot, \cdot \rangle$  is any inner product on  $\mathbf{E}$  with respect to which  $\mathbf{E}_+ = \mathbf{E}_+^\perp$ , one can show that  $O(\mathbf{E}) \cap \text{Aut}(\mathbf{E})$  is the stabilizer of some  $a \in \mathbf{E}_+^\circ$  ([15], Proposition I.1.9). It follows that  $O(\mathbf{E}) \cap G \leq G_a$ . Since  $G$  acts transitively on  $\mathbf{E}_+^\circ$ , we can find some  $g \in G$  with  $ga = u$ ; replacing  $\langle \cdot, \cdot \rangle$ , if necessary, by the inner product  $\langle x, y \rangle_g := \langle gx, gy \rangle$  — which is also self-dualizing, by [15], Proposition I.1.7 — we can assume that  $a = u$ , whence, that  $O(\mathbf{E}) \cap G \leq G_u$ .  $\square$

This gives us part (a) of the Theorem.

Now let  $K = G \cap O(\mathbf{E})$ , and let  $\mathfrak{k}$  denote  $K$ 's Lie algebra. Notice that  $\mathfrak{k} = \mathfrak{g} \cap \mathfrak{o}(\mathbf{E}) = \{X \in \mathfrak{g} \mid X^\dagger = -X\}$ .

**Corollary 1:** Let the inner product on  $\mathbf{E}$  be chosen as per Lemma C. Let  $K = G \cap O(\mathbf{E})$ , and let  $\mathfrak{k}$  be the Lie algebra of  $K$ , and let  $\mathfrak{g}_u$  denote the Lie algebra of  $G_u \leq G$ . Then

- (a)  $G_u$  is connected;
- (b)  $G_u = K$ ;
- (c)  $\mathfrak{g}_u = \mathfrak{k}$ ;

*Proof:* (a)  $G/G_u$  is homeomorphic to the simply connected space  $\mathbf{E}_+^\circ$ ; hence, as  $G$  is connected, so is  $G_u$  (see, e.g., [23], Proposition 1.94). It follows that if  $G_u$  and  $K$  have the same Lieq algebra, they coincide — in other words, (b) follows from (c). To prove (c), note that since  $K \leq G_u$ , by the choice of inner product, we have  $\mathfrak{k} \leq \mathfrak{g}_u$ . For every  $X \in \mathfrak{g}_u \leq \mathfrak{g}$ , we have the decomposition  $X = X_1 + X_2$  with  $X_1$  self-adjoint and  $X_2$ , skew-adjoint. Since  $X_2 \in \mathfrak{k} \subseteq \mathfrak{g}_u$ , it follows that  $X_1 = X - X_2 \in \mathfrak{g}_u$  as well, whence,  $e^{tX_1} \in G_u$  for all  $t$ . However,  $G_u$  is compact, so this implies that  $e^{tX_1}$  is bounded as a function of  $t$ . Since  $X_1$  is self-adjoint, this is possible only if  $X_1 = 0$ . Hence,  $X = X_2 \in \mathfrak{k}$ , and we have  $\mathfrak{g}_u \leq \mathfrak{k}$ .  $\square$

*Proof of Part (b):* Suppose now that  $G$  is self-adjoint with respect to the self-dualizing inner product chosen in Lemma C. Then, for every  $X \in \mathfrak{g}$ ,  $X^* \in \mathfrak{g}$ . To see this, let  $X = \gamma'(0)$  where  $\gamma : \mathbb{R} \rightarrow G$  is a smooth path with  $\gamma(0) = \mathbf{1}$ , and note that  $\gamma^* : t \mapsto \gamma(t)^*$  is another such path, with  $\gamma^{*'}(0) = X^*$ . It now follows that, for every  $X \in \mathfrak{g}$ ,  $X_1 := (X + X^*)/2$  and  $X_2 = (X - X^*)/2$  also lie in  $\mathfrak{g}$ ; we have  $X = X_1 + X_2$ , so that  $\mathfrak{g}$  decomposes as the direct sum  $\mathfrak{g} = \mathfrak{p} + \mathfrak{k}$ , where  $\mathfrak{p}$  is the space of self-adjoint elements of  $\mathfrak{g}$  and  $\mathfrak{k}$ , the space of skew-adjoint elements. This establishes part (b) of the Theorem.

*Remark:* Note also, for later reference, that if  $X, Y \in \mathfrak{p}$ , we also have  $[X, Y] \in \mathfrak{g}$  and  $[X, Y]^\dagger = [Y, X] = -[X, Y]$ , i.e.,  $[X, Y] \in \mathfrak{k}$ .

The interior,  $\mathbf{E}_+^\circ$ , of the cone  $\mathbf{E}_+$  is a smooth manifold, on which  $G$  acts smoothly. Thus, we have a canonical smooth mapping  $\phi : G \rightarrow \mathbf{E}_+^\circ$  given by  $g \mapsto gu$ . Differentiating this, we obtain a linear mapping  $d\phi(\mathbf{1}) : \mathfrak{g} \rightarrow T_u(\mathbf{E}_+^\circ) = \mathbf{E}$ . Explicitly, if  $X = \gamma'(0) \in \mathfrak{g}$ , where  $\gamma$  is a smooth path in  $G$  with  $\gamma(0) = \mathbf{1}$ , then

$$d\phi(\mathbf{1})(X) = \frac{d}{dt}\phi(\gamma(t))_{t=0} = \gamma'(0)u = Xu. \quad (3)$$

**Lemma D:** Let  $X \in \mathfrak{g}$ . Then  $X \in \mathfrak{k}$  iff  $Xu = 0$ .

*Proof:* Suppose  $d\phi(\mathbf{1})(X) = Xu = 0$ , where  $X = \gamma'(0)$ . The vector-valued function  $v(t) = e^{tX}u$  then satisfies

$$v' = Xv = Xe^{tX}u = e^{tX}Xu = 0.$$

It follows that  $v$  is constant, i.e., that  $e^{tX}u = u$  for all  $t$ . But then  $e^{tX} \in G_u$ , so that  $X = \frac{d}{dt}e^{tX}|_{t=0} \in \mathfrak{k}$ . Conversely, if  $X \in \mathfrak{k}$ , then  $X = \gamma'(0)$  where  $\gamma(t) \in K$  and  $\gamma(0) = \mathbf{1}$ , so that  $Xu = \gamma'(0)u = [\frac{d}{dt}(\gamma(t)u)]_{t=0} - [\gamma(t)\frac{d}{dt}u]_{t=0} = [\frac{d}{dt}(\gamma(t)u)]_{t=0} = \frac{d}{dt}u|_{t=0} = 0$  (the last equality uses  $\gamma(t) \in K$  and  $Ku = u$ ).  $\square$

**Corollary 2:**  $d\phi(\mathbf{1}) : \mathfrak{p} \simeq \text{ran}(d\phi(\mathbf{1})) \leq \mathbf{E}$ .

*Proof:* By Lemma D and (3),  $\mathfrak{k}$  is the kernel of  $d\phi(\mathbf{1})$ ; as established above (in the proof of part (b) of Lemma C,  $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ .  $\square$

**Lemma E:**  $\text{ran}(d\phi(\mathbf{1}))$  is all of  $\mathbf{E}$ , i.e.,  $d\phi(\mathbf{1}) : \mathfrak{p} \simeq \mathbf{E}$ .

*Proof:* Note, first, that  $T_g(G) = gT_1(G) = g\mathfrak{g}$  for any  $g \in G$ . We also have

$$(d\phi(g))(gX) = gXu$$

for all  $X \in \mathfrak{g}$ . Hence,  $\text{ran}(d\phi(g)) = \text{gran}(d\phi(\mathbf{1}))$ . If the latter is not all of  $\mathbf{E}$ , then every  $g$  is a critical point of  $\phi$ , whence, every point  $\phi(g) = gu \in \mathbf{E}$  is a critical value. Sard's Theorem now tells us that  $G_u = \mathbf{E}_+^\circ$  has measure zero, a contradiction.  $\square$

*Construction of the Jordan product* Now define  $L_x \in \mathfrak{p}$  to be the unique self-adjoint element of  $\mathfrak{g}$  with  $L_x u = x$ . Set  $x \bullet y = L_x y$  for all  $x, y \in \mathbf{E}$ . This is evidently bilinear. A series of computations (see [15], pp. 49-50) shows that it makes  $\mathbf{E}$  a formally real Jordan algebra with identity element  $u$ . Specifically,

(1)  $x \bullet y = y \bullet x$  since  $x \bullet y - y \bullet x = x \bullet (y \bullet u)u - y \bullet (x \bullet u) = [L_x, L_y]u = 0$  (as remarked above following the proof of Corollary 1,  $X, Y \in \mathfrak{p} \Rightarrow [X, Y] \in \mathfrak{k}$ , and, by Lemma D,  $\mathfrak{k}u = 0$ );

(2)  $x \bullet u = L_x u = x$  by definition of  $L_x$ , so  $u$  serves as the identity;

(3) The product satisfies the Jacobi identity. This is proved exactly as in [15]. Note that the argument uses the fact that the inner product is associative, which follows from  $L_x$  being self-adjoint. This also then gives us that the Jordan algebra  $\mathbf{E}$  is formally real.

This completes the proof of the Koecher-Vinberg Theorem.  $\square$